

Project 1

Philosophy Data EDA: How do we recommend a philosopher/school to a rookie

If you were asked by a friend about how to start their journey to philosophy, how would you react? Talking about the complex doctrine and concepts of each famous philosopher for half an hour may not help a rookie find out a suitable startpoint, since they may not fully understand what you are talking about. Today, I am trying to deal with this issue with the help of data.

First, let us set up the environment by importing some necessary packages and loading our data.

Viewing the basic information and checking for missing data

The information of the data:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 360808 entries, 0 to 360807
Data columns (total 12 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   title            360808 non-null   object  
 1   author           360808 non-null   object  
 2   school           360808 non-null   object  
 3   sentence_spacy   360808 non-null   object  
 4   sentence_str     360808 non-null   object  
 5   original_publication_date  360808 non-null   int64  
 6   corpus_edition_date  360808 non-null   int64  
 7   sentence_length   360808 non-null   int64  
 8   sentence_lowered  360808 non-null   object  
 9   tokenized_txt    360808 non-null   object  
 10  lemmatized_str   360808 non-null   object  
 11  num_of_tokens    360808 non-null   int64  
dtypes: int64(4), object(8)
memory usage: 33.0+ MB
None
```

The shape of the raw data:
(360808, 12)

Describing table:

	original_publication_date	corpus_edition_date	sentence_length	\
count	360808.000000	360808.000000	360808.000000	
mean	1326.800908	1995.155642	150.790964	
std	951.492193	23.002287	104.822072	
min	-350.000000	1887.000000	20.000000	
25%	1641.000000	1991.000000	75.000000	
50%	1817.000000	2001.000000	127.000000	
75%	1949.000000	2007.000000	199.000000	
max	1985.000000	2016.000000	2649.000000	

	num_of_tokens
count	360808.000000
mean	25.693216
std	17.766261
min	0.000000
25%	13.000000
50%	22.000000
75%	34.000000
max	398.000000

Checking missing data:

title	0
author	0
school	0
sentence_spacy	0
sentence_str	0
original_publication_date	0
corpus_edition_date	0
sentence_length	0
sentence_lowered	0
tokenized_txt	0
lemmatized_str	0
num_of_tokens	0

dtype: int64

```
      title author school \
0  Plato - Complete Works  Plato  plato

                           sentence_spacy \
0  What's new, Socrates, to make you leave your ...

                           sentence_str \
0  What's new, Socrates, to make you leave your ...

original_publication_date  corpus_edition_date  sentence_length \
0            -350                  1997                 125

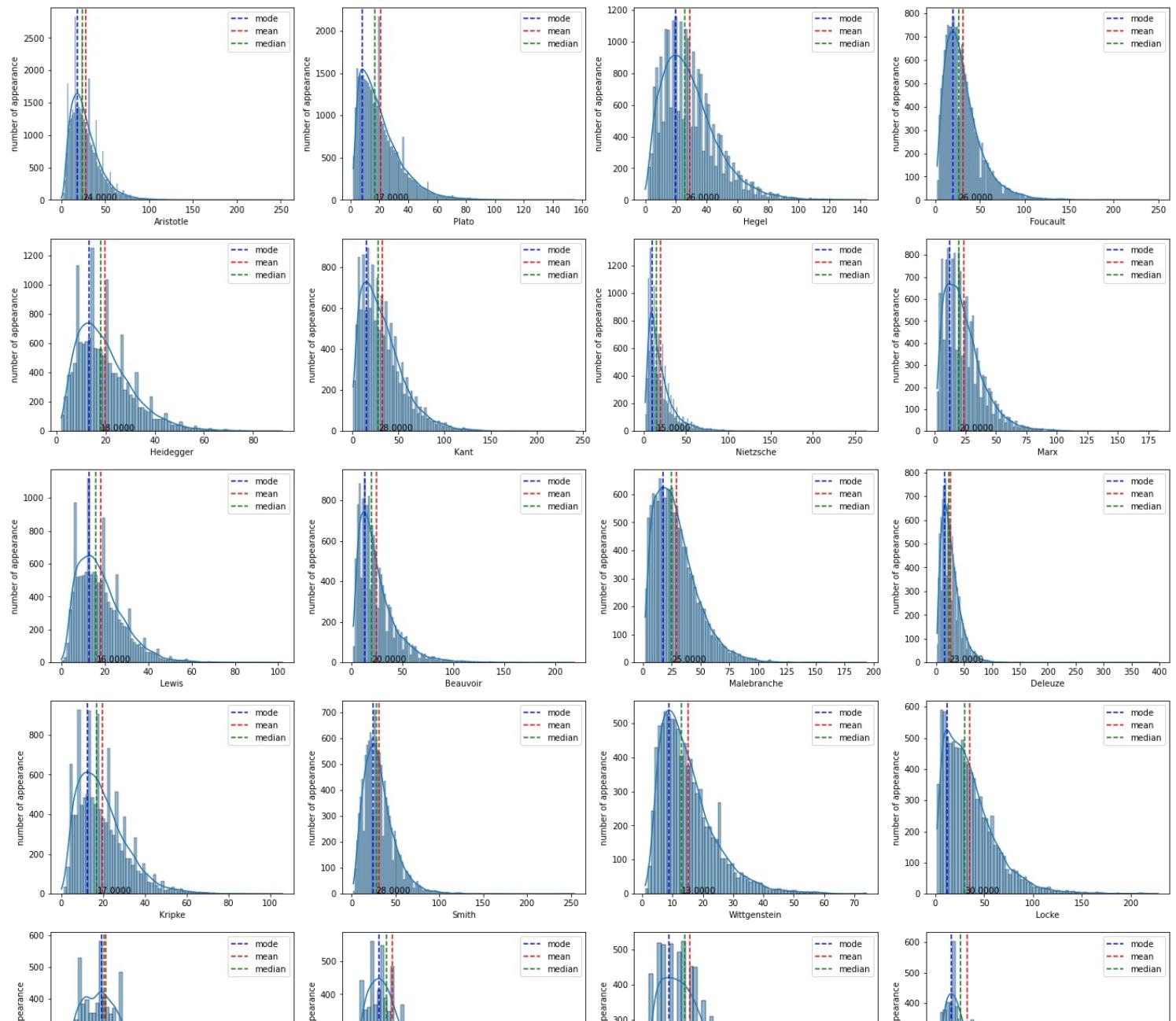
                           sentence_lowered \
0  what's new, socrates, to make you leave your ...

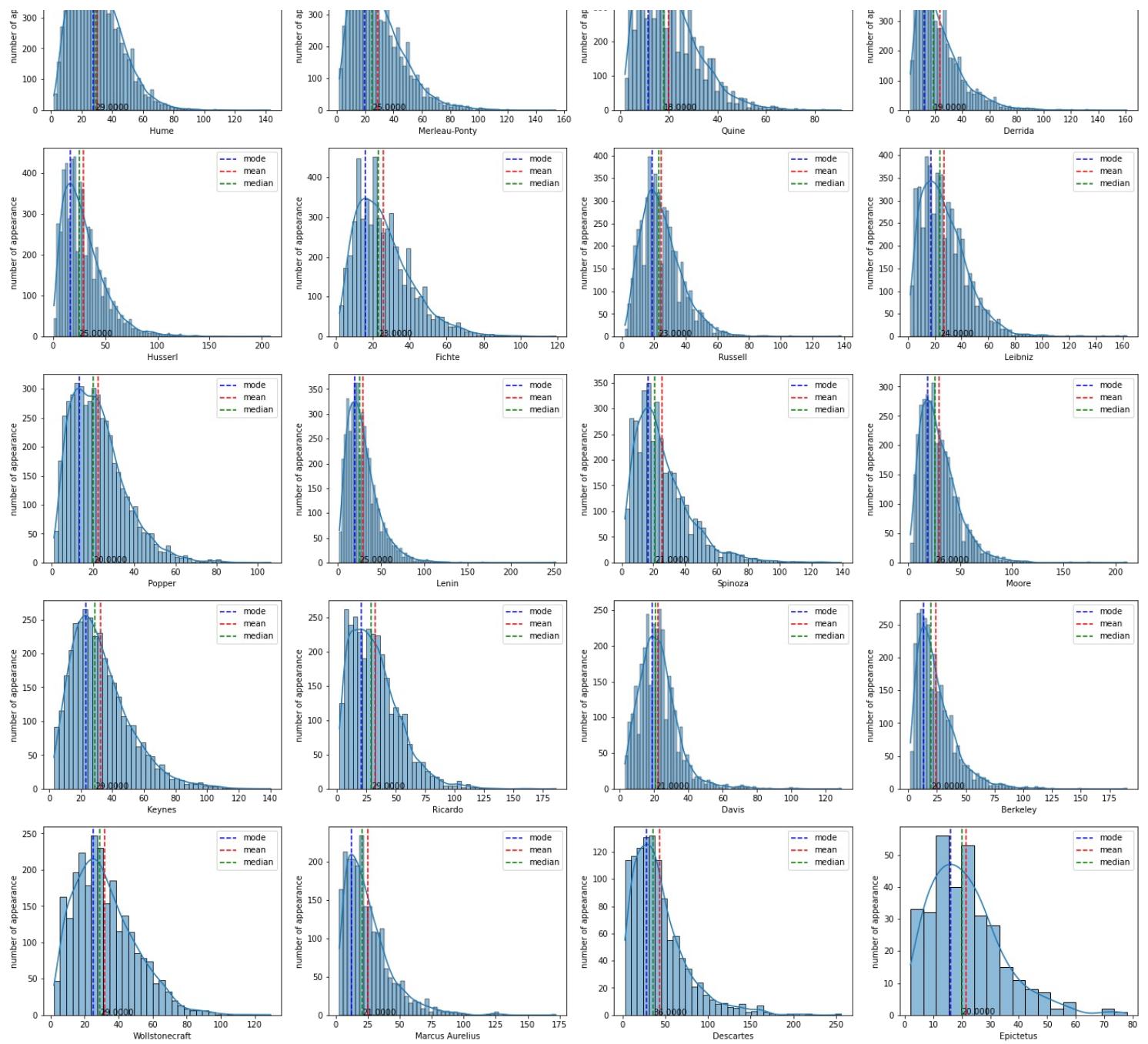
                           tokenized_txt \
0  ['what', 'new', 'socrates', 'to', 'make', 'you...

                           lemmatized_str  num_of_tokens
0  what be new , Socrates , to make -PRON- lea...           23
```

Everything seems all good, and now we can start some analysis.

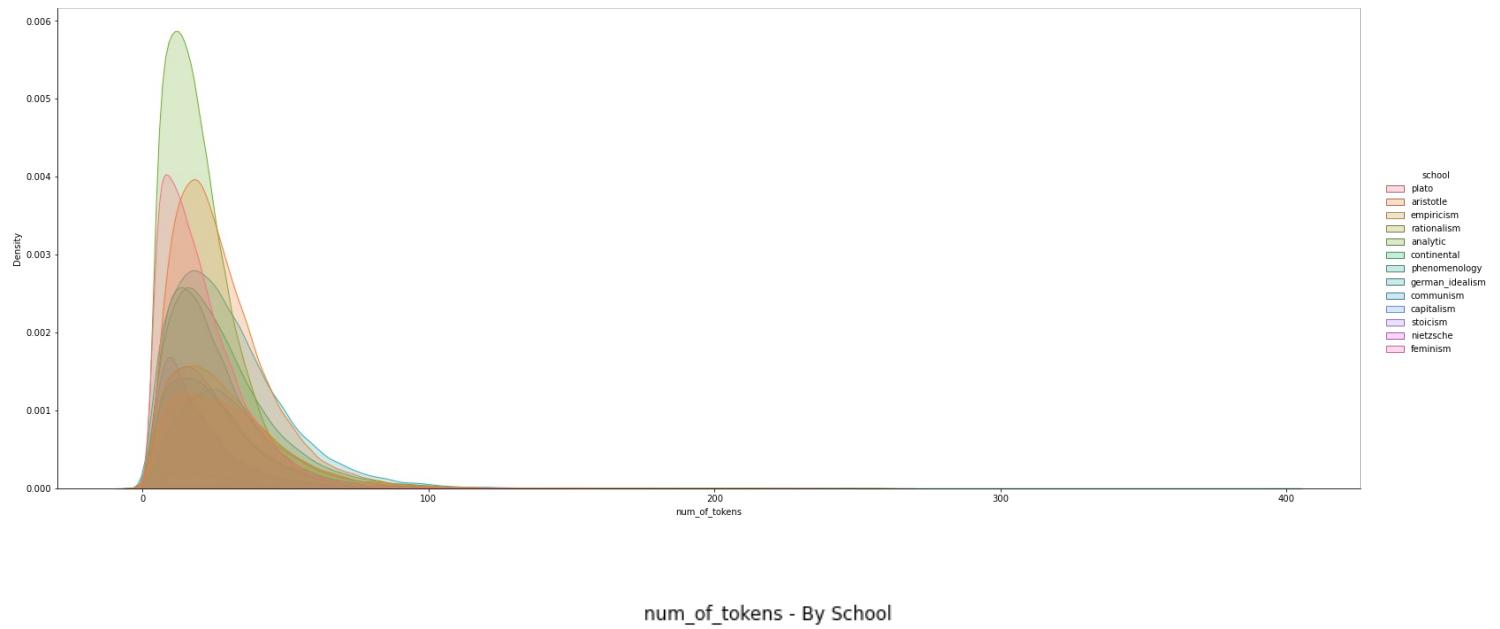
The first feature is the number of tokens of sentences by philosophers, which indicates the difficulties of understanding the content. After all, not anyone's plan is to being professional because they may not have enough time for consuming, and a book full of complex sentences is not an ideal choice for a beginner. Therefore, here are graphs showing the hardness of reading according to the length of sentences.

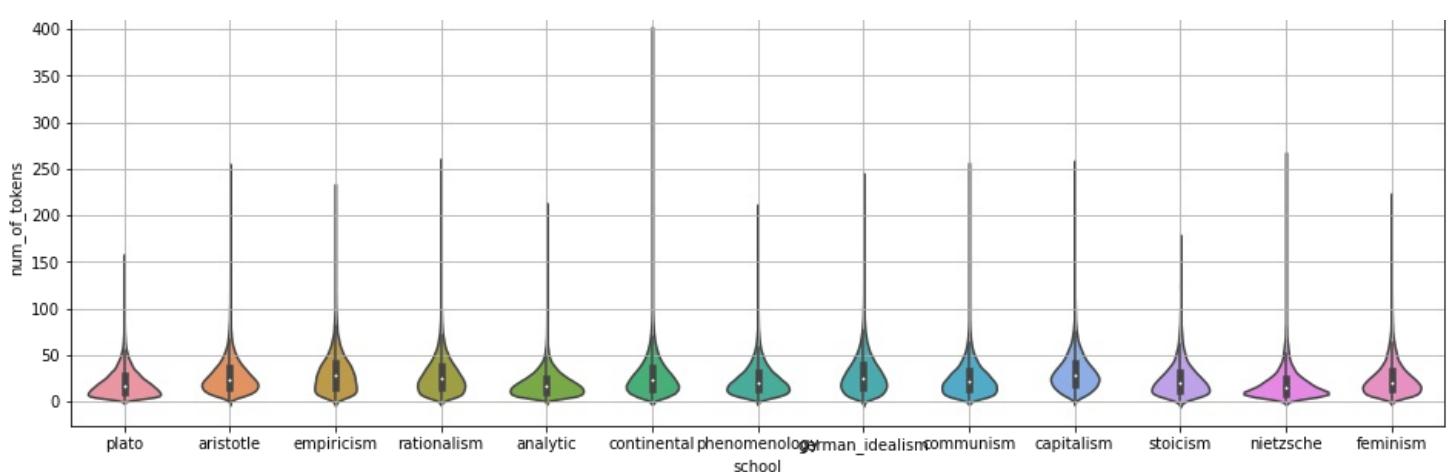




As we can see, Plato has relatively shorter sentences comparing to others', therefore he might be a good choice for the beginners

<seaborn.axisgrid.FacetGrid at 0x1f18aa43ee0>

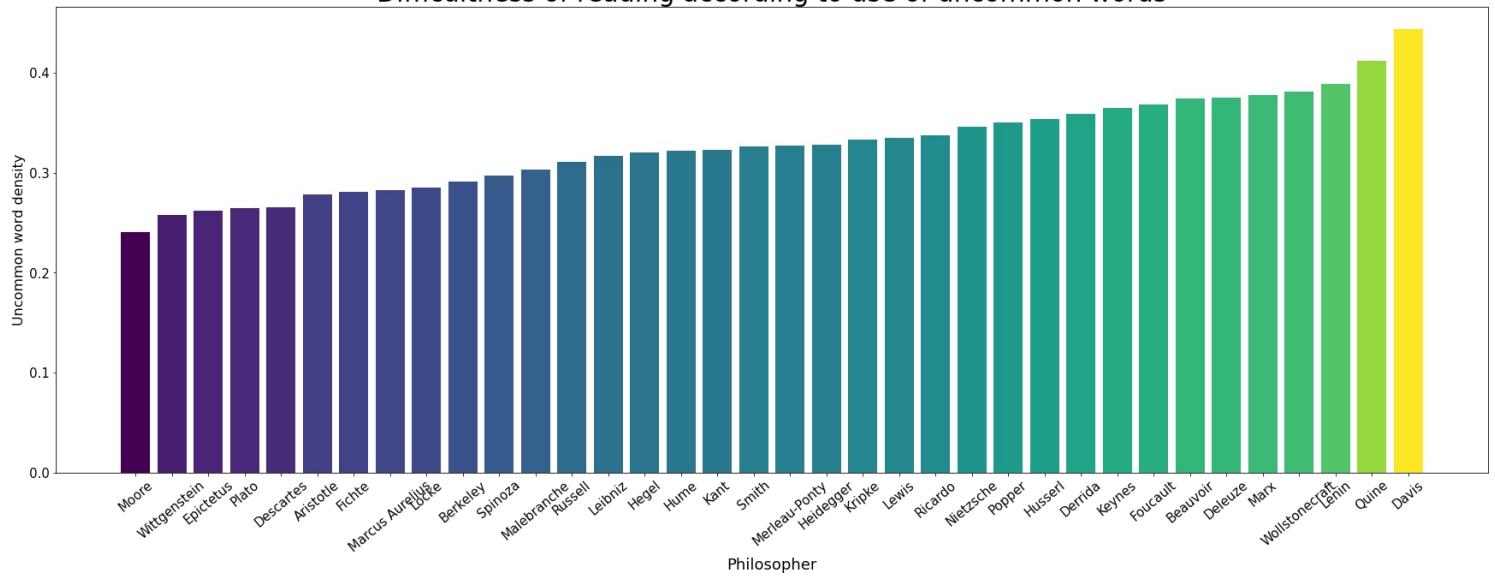




Besides, we can also use the density of uncommon words as criterion of difficultness of reading. Here are the graphs showing the comparison of this feature of respective philosophers and schools.

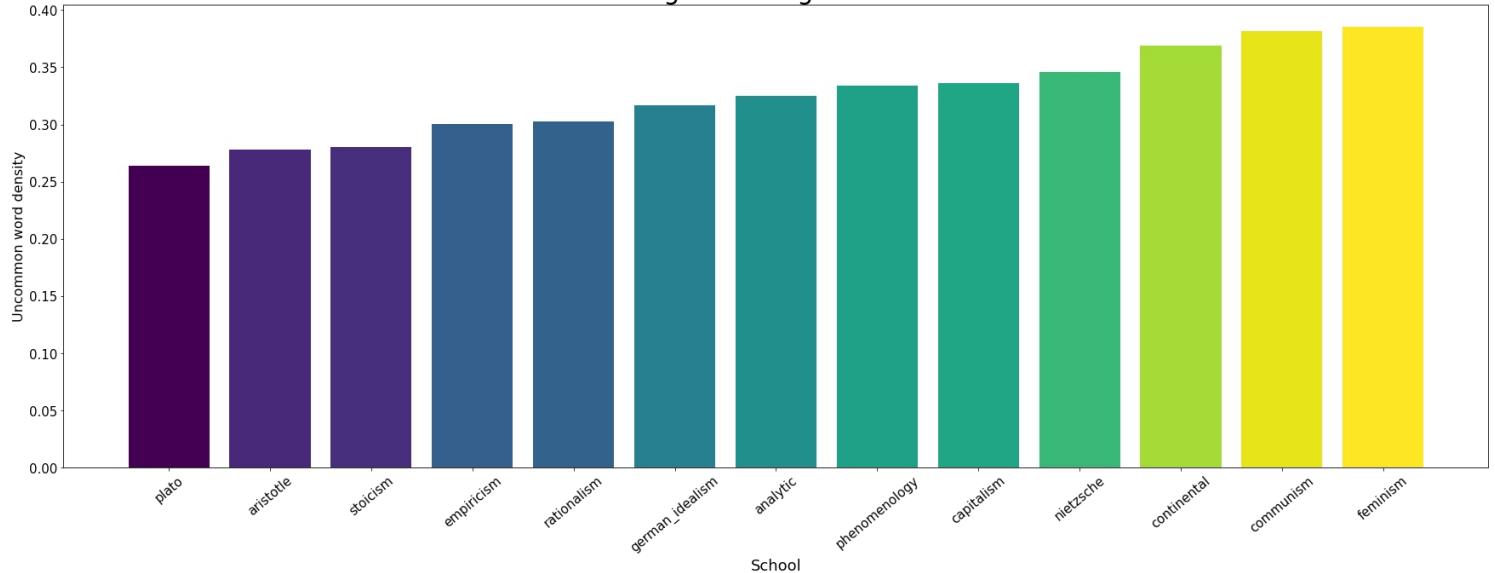
Text(0.5, 1.0, 'Difficultness of reading according to use of uncommon words')

Difficulty of reading according to use of uncommon words



Text(0.5, 1.0, 'Difficultness of reading according to use of uncommon words')

Difficulty of reading according to use of uncommon words



Analogously, schools and philosophers like Plato and Aristotle might be more friendly for beginners, for they using less uncommon words.

Except for hardness of reading, wordclouds and sentiment analysis are easy and straightforward tools to help us get a summary picture about the focus and emotion of different schools and philosophers, which can enable us to take personal preference into consideration.

School = ANALYTIC :



School = ARISTOTLE :



School = GERMAN IDEALISM

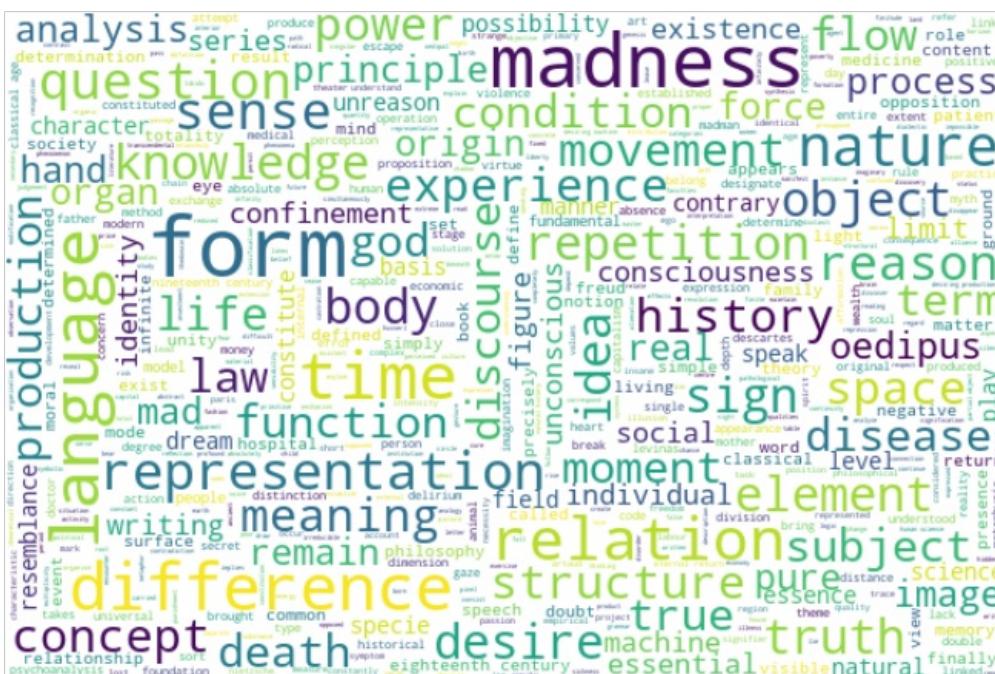




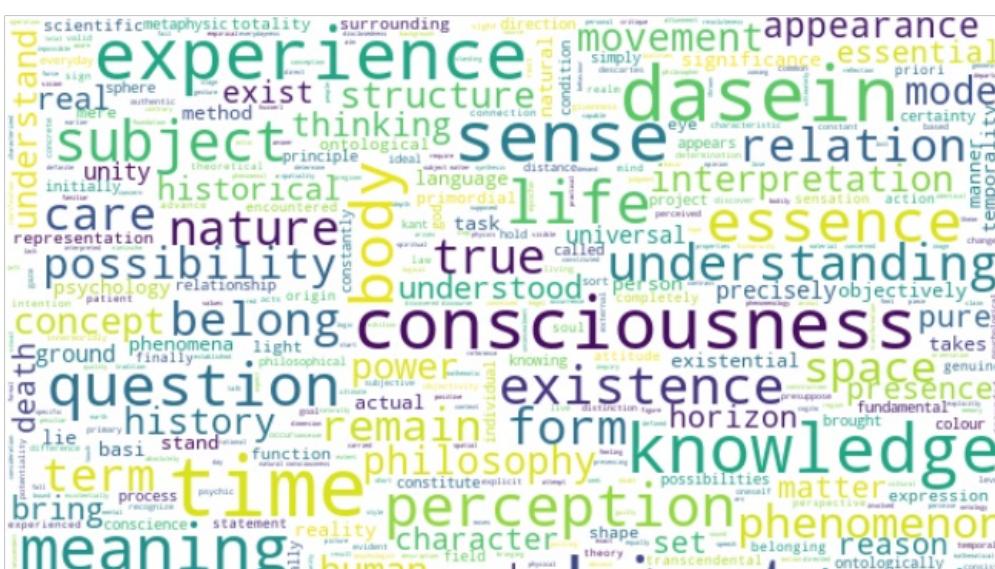
School = PLATO :



School = CONTINENTAL



School = PHENOMENOLOGY





School = RATIONALISM :



School = EMPIRICISM :



School = FEMINISM :





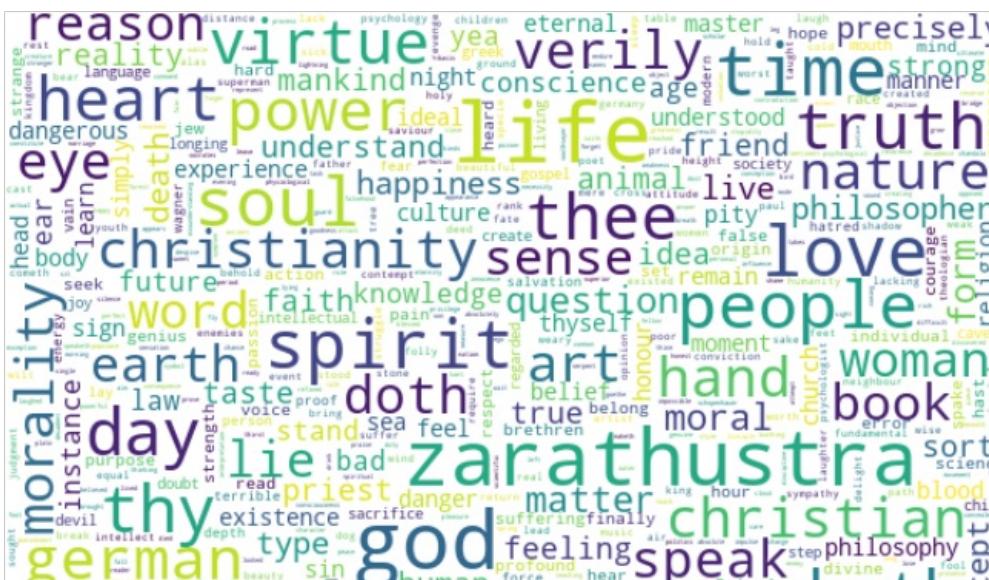
School = CAPITALISM



School = COMMUNISM



School = NIETZSCHE



view surf pure short mountain values war deep wisdom hitherto learned noble meaning saint light hath
 history contrary secret desire instinct called regard conc nowaday
 values natural wisdom freedom desire recondition doctrine heaven european account born

School = STOICISM :

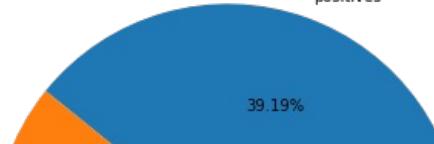


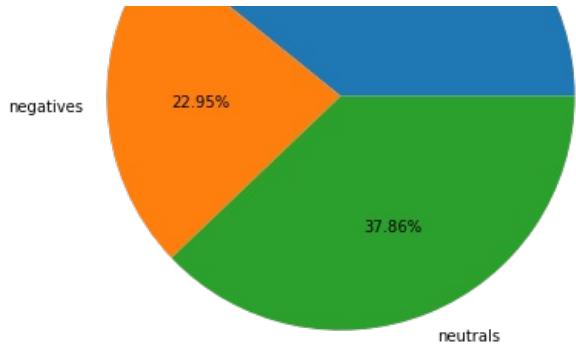
Elapsed time: 32.35

```
[nltk_data] Downloading package vader_lexicon to
[nltk_data]   C:\Users\shy\AppData\Roaming\nltk_data...
```

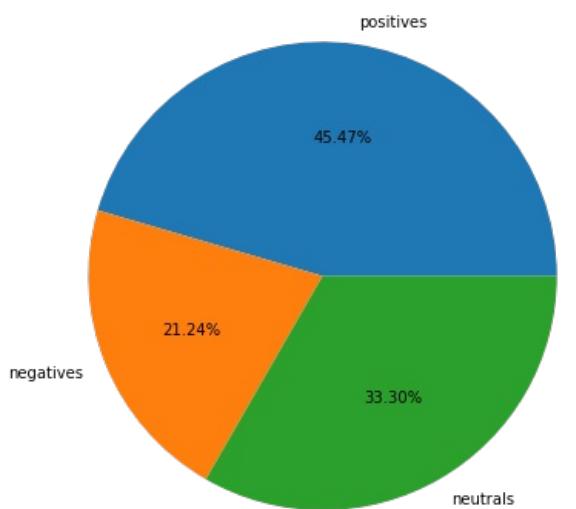
C:\Users\shy\AppData\Local\Temp\ipykernel_10508\1045417455.py:38: RuntimeWarning: More than 20 figures have been opened. Figures created through the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly closed and may consume too much memory. (To control this warning, see the rcParam `figure.max_open_warning`).
plt.figure(figsize = (7, 7))

Sentiment Analysis for Philosopher: Aristotle

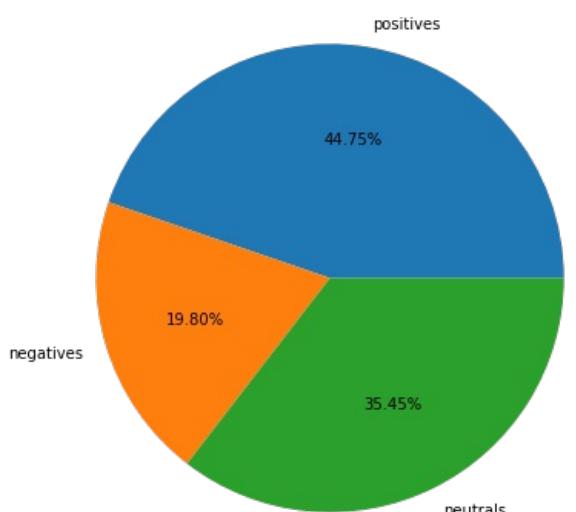




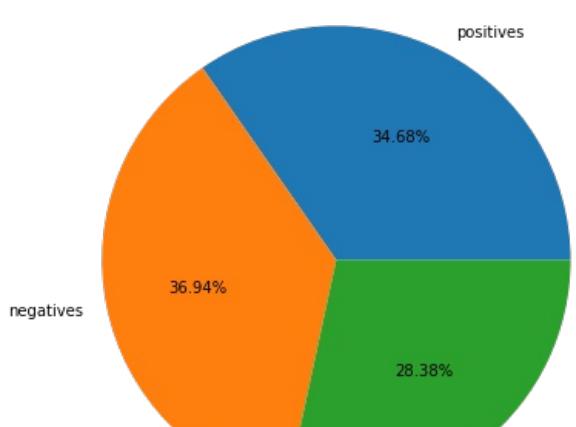
Sentiment Analysis for Philosopher: Plato



Sentiment Analysis for Philosopher: Hegel

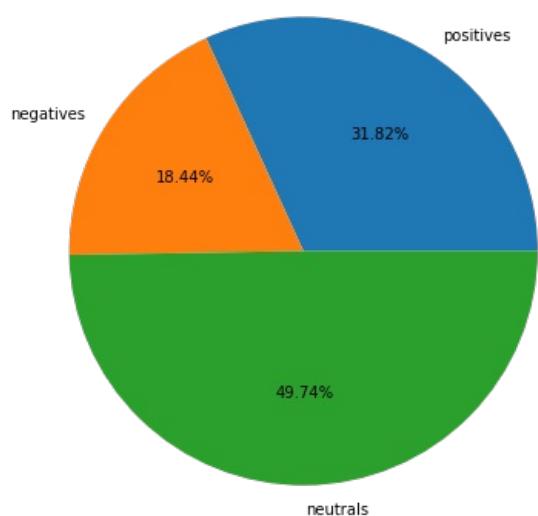


Sentiment Analysis for Philosopher: Foucault

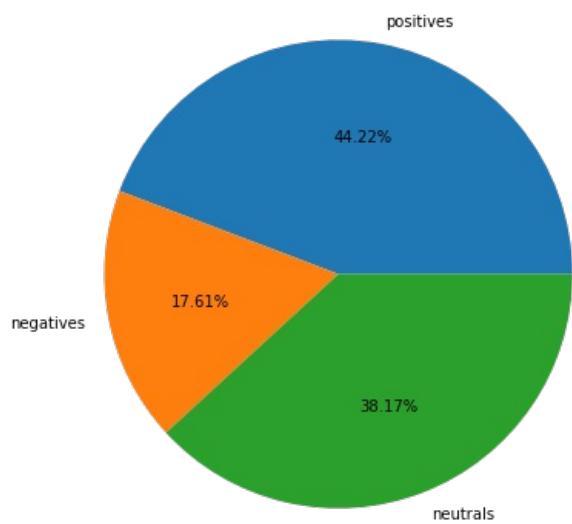




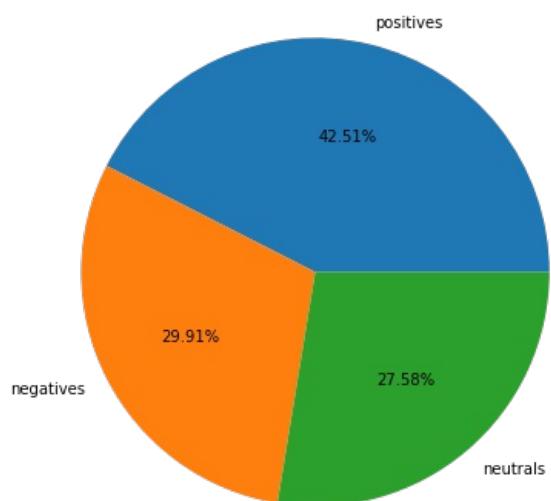
Sentiment Analysis for Philosopher: Heidegger



Sentiment Analysis for Philosopher: Kant

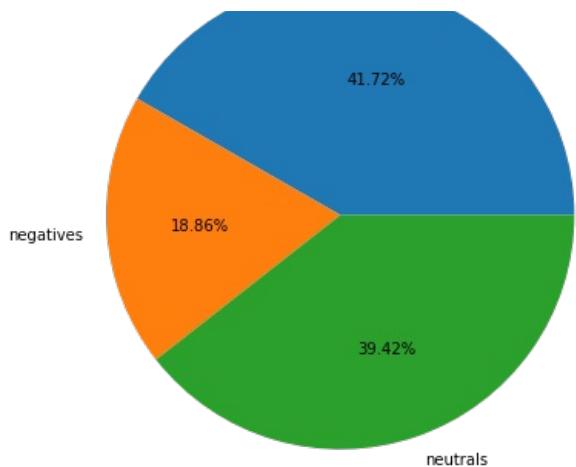


Sentiment Analysis for Philosopher: Nietzsche

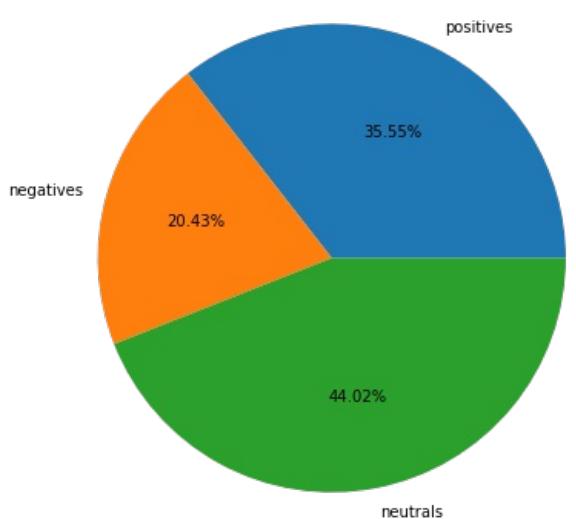


Sentiment Analysis for Philosopher: Marx

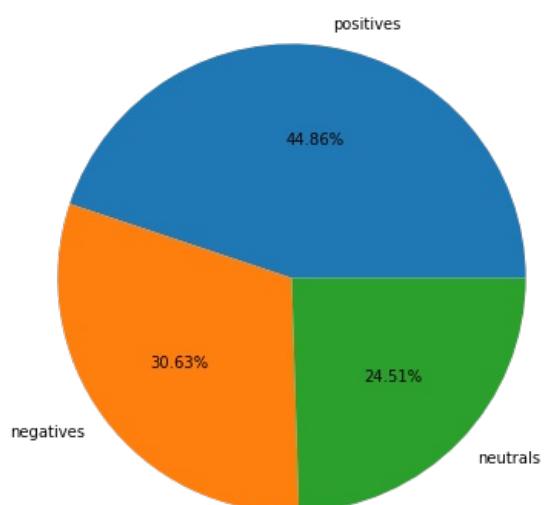




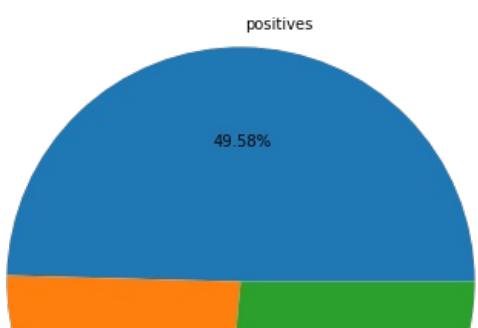
Sentiment Analysis for Philosopher: Lewis

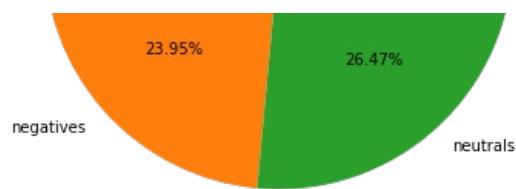


Sentiment Analysis for Philosopher: Beauvoir

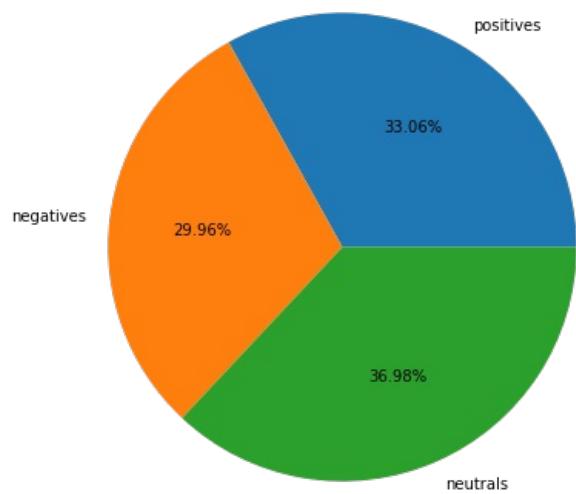


Sentiment Analysis for Philosopher: Malebranche

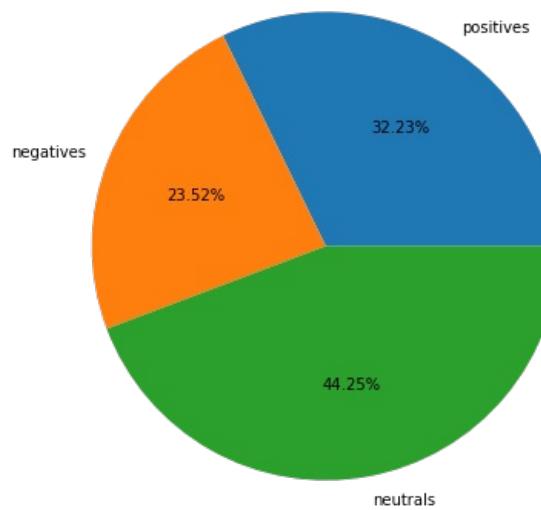




Sentiment Analysis for Philosopher: Deleuze

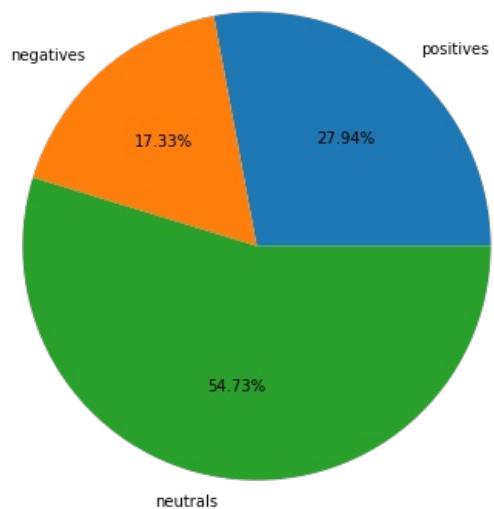


Sentiment Analysis for Philosopher: Kripke

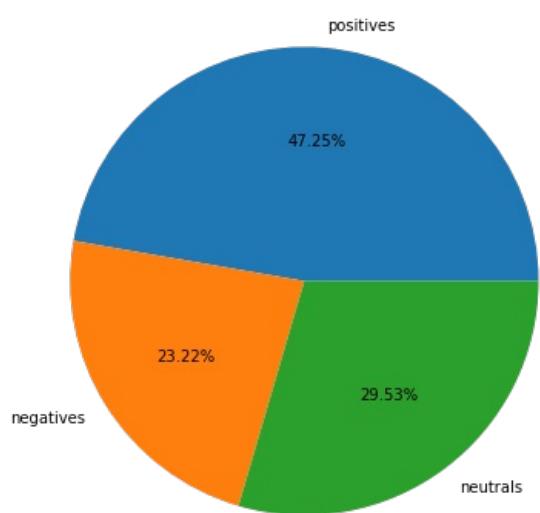


Sentiment Analysis for Philosopher: Smith

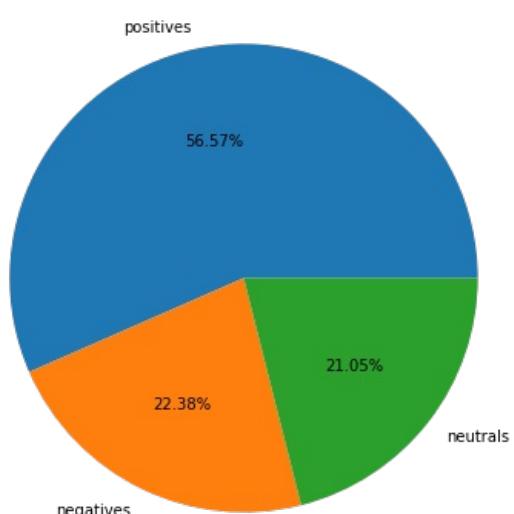
Sentiment Analysis for Philosopher: Wittgenstein



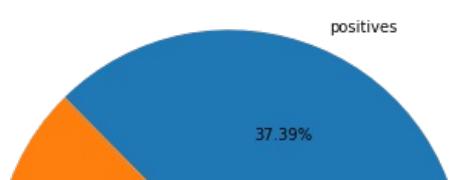
Sentiment Analysis for Philosopher: Locke

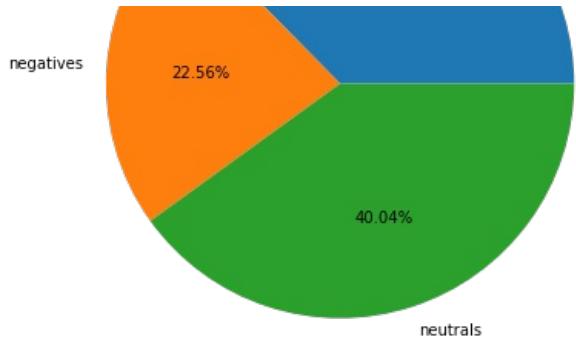


Sentiment Analysis for Philosopher: Hume

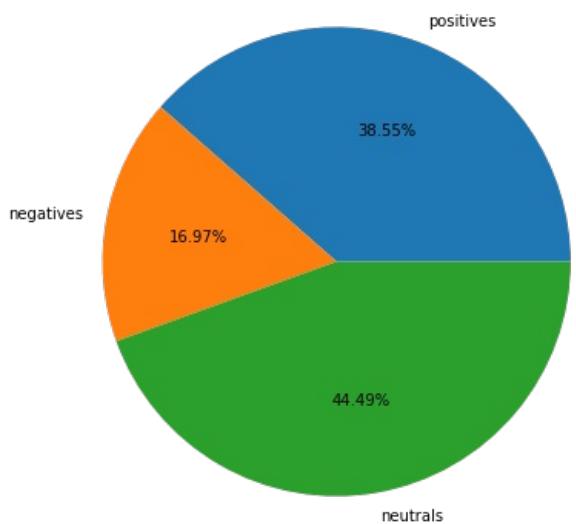


Sentiment Analysis for Philosopher: Merleau-Ponty

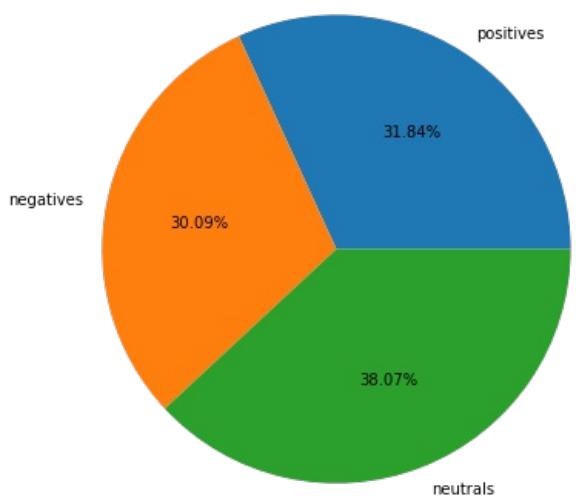




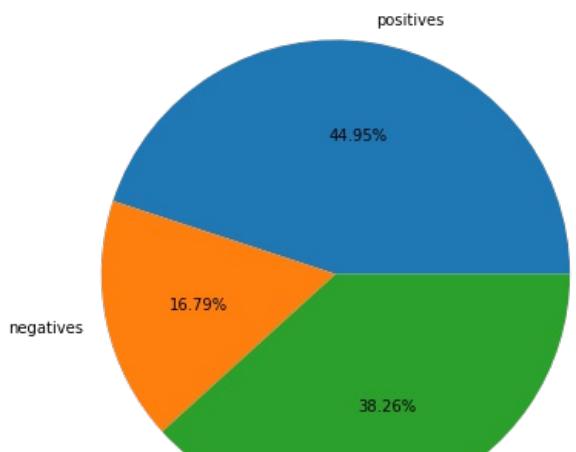
Sentiment Analysis for Philosopher: Quine



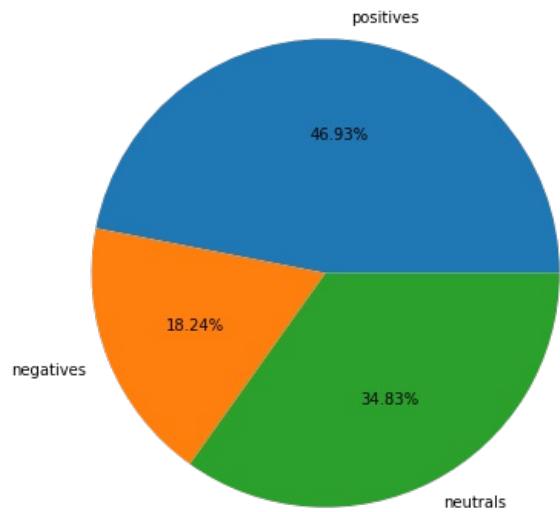
Sentiment Analysis for Philosopher: Derrida



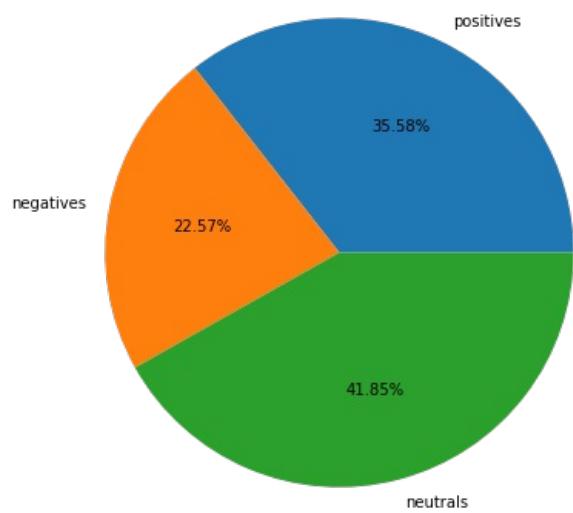
Sentiment Analysis for Philosopher: Husserl



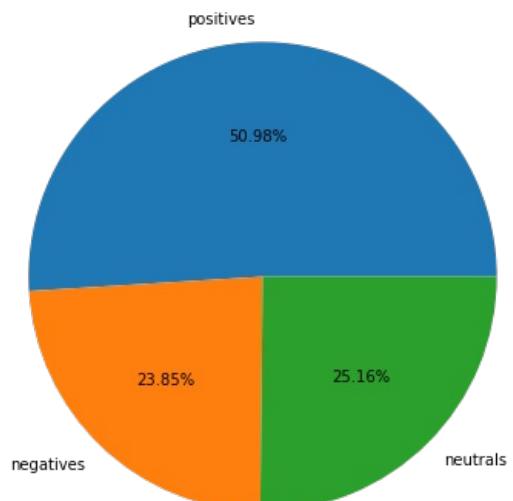
Sentiment Analysis for Philosopher: Fichte



Sentiment Analysis for Philosopher: Russell

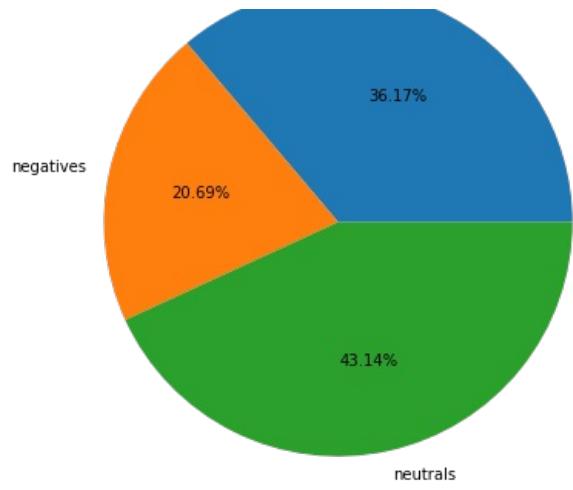


Sentiment Analysis for Philosopher: Leibniz

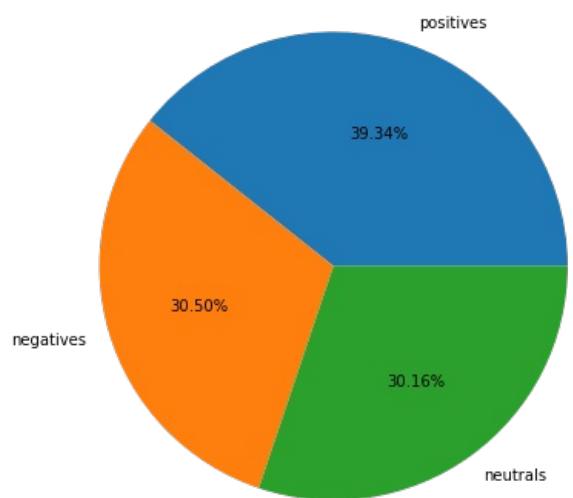


Sentiment Analysis for Philosopher: Popper

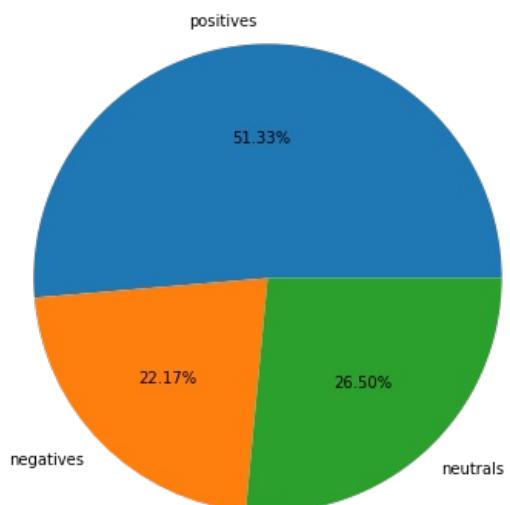




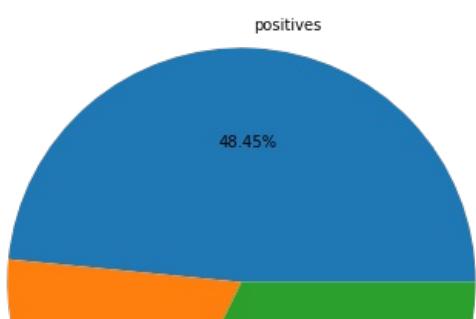
Sentiment Analysis for Philosopher: Lenin

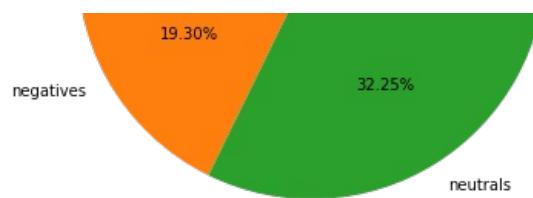


Sentiment Analysis for Philosopher: Spinoza

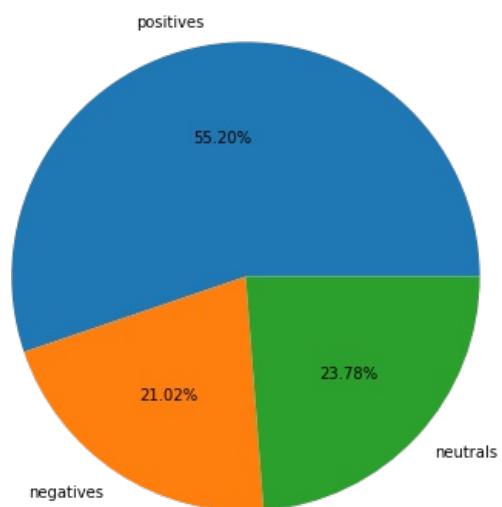


Sentiment Analysis for Philosopher: Moore

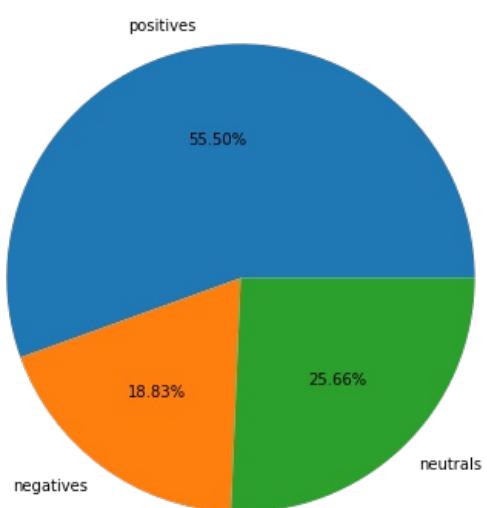




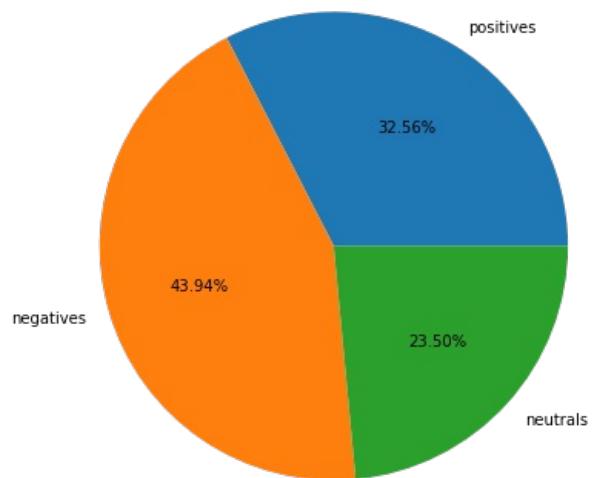
Sentiment Analysis for Philosopher: Keynes



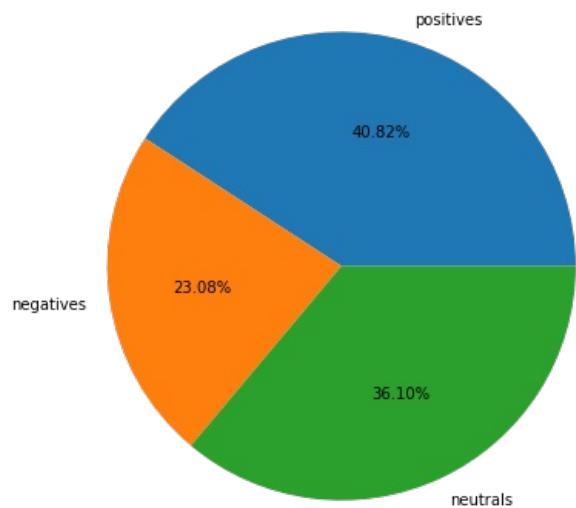
Sentiment Analysis for Philosopher: Ricardo



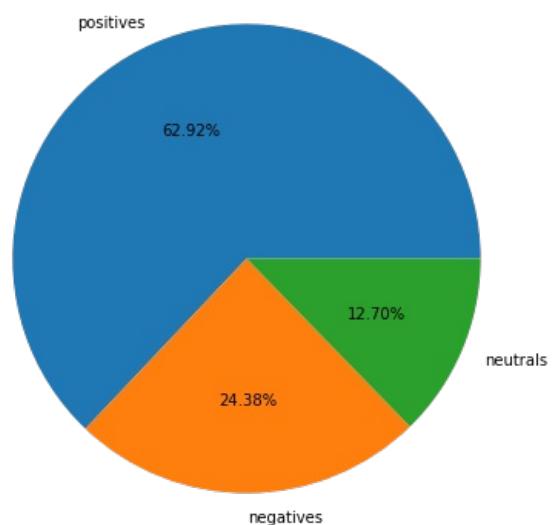
Sentiment Analysis for Philosopher: Davis



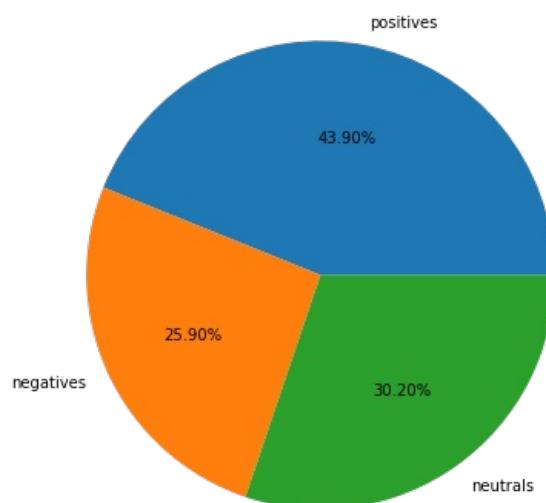
Sentiment Analysis for Philosopher: Berkeley



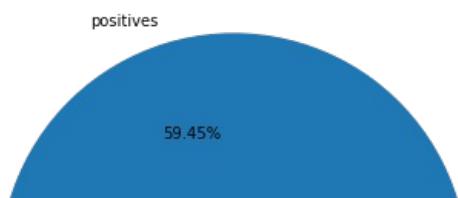
Sentiment Analysis for Philosopher: Wollstonecraft

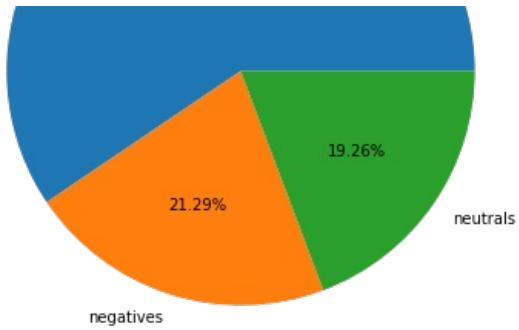


Sentiment Analysis for Philosopher: Marcus Aurelius

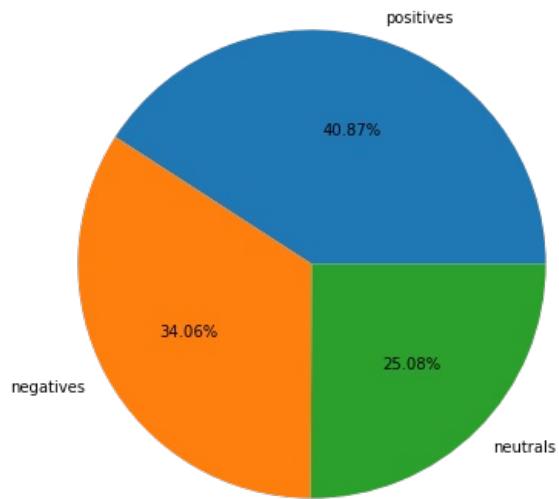


Sentiment Analysis for Philosopher: Descartes





Sentiment Analysis for Philosopher: Epictetus



With help of data, we are now able to recommend a suitable startpoint and even a direction of further study to our friends who want to dive into the ocean of philosophy

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js